AD 154613

Armed Services Technical Information Agency

ARLINGTON HALL STATION
ARLINGTON 12 VIRGINIA

FOR MICRO CARD CONTROL ONLY

1 OF 1

Research Report

154 613

Research Project NM 18 02 99 Subtask 1 Report No. 74

DELAYED RESPONSE: EFFECTS UPON SPEECH RECEPTION AND SPEAKER INTELLIGIBILITY

FC

U. S. NAVAL SCHOOL OF AVIATION MEDICINE

U. S. NAVAL AIR STATION PENSACOLA, FLORIDA

U. S. NAVAL SCHOOL OF AVIATION MEDICINE
NAVAL AIR STATION
PENSACOLA, FLORIDA

JOINT PROJECT

The Ohio State University Research Foundation
Columbus, Ohio, under Contract N6ONR 22525
Office of Naval Research Project Designation No. NR 145-993

and

U. S. Naval School of Aviation Medicine
The Bureau of Medicine and Surgery Project NM 18 02 99
Subtask 1 Report No. 74

DELAYED RESPONSE: EFFECTS UPON SPEECH RECEPTION AND SPEAKER INTELLIGIBILITY

Report by

Gilbert C. Tolhurst

Approved by

John W. Black
Project Director
and
Captain Ashton Graybiel, MC, USN
Director of Research

Released by

Captain Langdon C. Newman, MC, USN
Commanding Officer

15 August 1957

SUMMARY PAGE

THE PROBLEM

Individuals who listened to multiple-choice tests and other individuals who repeating the test-word groups after hearing them aurally, acted as speakers, were instructed to delay their responses to the word group items until a response signal tone was presented. The tone followed the last word in the three word groupings by 0, 1, 2, 3, 4, or 5 seconds, randomly.

Other groups of individuals acted as listeners and still others as talkers under two specific response delay time conditions of either one or five seconds. The verbal materials were the PB write-down word lists.

FINDINGS

1. Increases in response delay times to multiple-choice test items resulted in incrementally higher scores of both listener reception and speaker intelligibility under the conditions of zero to five seconds response delay. Delays of five seconds yielded the highest scores.

2. When PB word lists were used as stimulus material under the conditions of one or five seconds delay, the speaker intelligibility scores showed an increase at the longer delay time, but the speech reception scores were reversed, the higher scores being associated with the shorter delay time.

INTRODUCTION

Several studies have been devoted to the exploration of the efficiency of certain methods by which a speaker acquires the verbal material he is to transmit (2,5, 7-11). The data obtained and the inferences drawn from the analyses assumed either minimal or no delay time between the time of reception and the retransmission of such material to listeners. Short time delays of 0.05, 0.08, and 0.09 second in a speaker's side-tone tend to improve his intelligibility (1), but the mechanism in side-tone modification is not strictly one of retransmission of speech. Another study, more clearly allied with response delay, found that delaying the onset of aural signals 0.23 second in relation to the visual facial gestures made no difference to the resultant intelligibility scores (12).

Other studies have had as parameters certain "stressful" situations or certain tasks inserted between the presentation of the stimulus word items and the listener responses to the words (4,6). The assumption seemed implicit that delaying the responses did not contribute to changes in speech reception efficiency.

The present study was designed to describe the function of response time delay of six incremental steps from zero to five seconds when the verbal material was the multiple-choice type tests used as speech reception tests and as speaker intelligibility tests. Additional data were obtained on two delay times using PB(Phonetically Balanced) write-down tests (3).

PROCEDURE

There were four sets of data obtained to test the hypotheses of no difference among the response delay times as they might have influenced: (a) the Speech Reception of multiple-choice test items, (b) the Speaker Intelligibility of the multiple-choice test items, (c) the Speech Reception of PB words, and (d) the Speaker Intelligibility of PB words.

MULTIPLE-CHOICE TESTS

(a) Speech Reception

Forms C and D of the multiple-choice tests were recorded by nine speakers. Each speaker read one three-word group item in each of the 24 lists comprising the two forms. The speakers were rotated so that the same speaker did not read the same three-word item in each list; however, no attempt was made to randomize the order in which speakers talked. This recording was dubbed so that a 1000 cps tonal burst followed the last word in each three-word item group by 0, 1, 2, 3, 4, or 5 seconds. The tonal "beep" delay remained the same for each list, but the delay times were distributed randomly among lists, except that each delay time occurred four times within the 24 lists.

Twelve panels of listeners, 15 to 24 per panel, were given routinized extemporaneous instructions as to how they were to respond to the test, with the additional admonition not to mark their papers until they heard the response tone following each word-group item. A short pre-test was administered prior to the administration of the test proper to acquaint the listeners with the method of test taking.

The stimulus tape playback was adjusted so that the speech levels averaged 80 db under the headset cushions (HS-33 headsets, PDR-3 earphones mounted in NAF-48490-1 doughnut cushions). The speech signals feeding the headset circuit were mixed, in line, with ASA white noise from an H. H. Scott generator, Model 811-A, at a +15 db signal/noise ratio. Playback of the speech was from an Ampex 600 magnetic tape recorder. The listening was done in a classroom situation having an average ambient noise level of 58 db re 0.0002 dyne/cm^2 , C scale of sound level meter.

(b) Speaker Intelligibility

Forty-eight speakers, four per panel, heard the identical nine-voiced delayed response Forms C and D, as did the listening panels in (a) above. Their task was to repeat the stimulus phrases after hearing the beep tone following each phrase. In effect, then, the result was to delay the speakers saying the stimulus phrases after receiving the "message" from zero to five seconds. Each speaker read six intelligibility lists, each list incorporating one of the delay times.

The speakers read the items from a small sound-treated room utilizing an Altec 21-C boom-mounted microphone attached to a standard HS-33 headset through which each speaker received the verbal material he was to repeat. The microphone was positioned along the cheek with the diaphragm of the microphone at the corner of the mouth; parallel incidence.

The speech signals were amplified and mixed with ASA white noise by an Altec 250-A console to deliver a +15 db signal/noise ratio. The combined signals and noise were then channeled to the headset circuit of panels of listeners in a larger sound room at an 80 db level under the earphone cushions. The responses of the listeners, 16 to 30 listeners per panel, to the multiple-choice tests yielded speaker intelligibility scores. The listeners received the same instructions as did those taking part in the speech reception portion of the study outlined in (a) above.

PB WORD TESTS

(c) Speech Reception

A single speaker recorded four of the 50-word PB lists using the same equipment described in part (b) of the present study. Two delay times, one and five seconds, were built into two recordings. For one of the recordings there was a 1000 cps beep tone presented following the stimulus word by one second and for the other two PB lists recording the tone followed the words by five seconds.

One-half of a panel of listeners (15 individuals) heard the PB lists with the one-second delay response and simultaneously the second one-half of the panel (15 other individuals) heard the tape with the five-second delay tone. Both groups were instructed to wait until they heard the beep tone following each word before they wrote down the word they heard. The voice signals were played back at a level of 80 db under the headset cushions. An ASA white noise was mixed with the voice signals at a +15 db signal/noise ratio and remained constant throughout the testing session.

(d) Speaker Intelligibility

Twenty-four individuals served as speakers in the PB intelligibility portion of the present study. Four speakers listened to a single-voiced recording of four PB lists (1-4), and their repetitions of the words were heard by a panel of listeners, one speaker per list. The one-voiced stimulus recording was the same as that heard by the listening panel in (c) above. Each speaker responded by repeating the stimulus words he heard both one and five seconds after he heard the word via his headset. The speakers heard the words at an 80 db SPL under the headset cushions mixed with ASA white noise set to yield a +15 db signal/noise ratio. The speakers were stationed in a small sound-proofed room.

The speaker's voice was picked up by an Altec 21-C condenser microphone boom-mounted on the headset, as in part (b) of the present study described above, and delivered to a panel of listeners (20 to 30 per panel). The voltages across the listeners' headsets were converted to sound-pressure level and averaged 80 db. The listening was done in "quiet", i.e., 42 db re 0.0002 dyne/cm^2 , C scale of the sound level meter. There were six replications of the above procedure.

RESULTS

The data obtained under the four testing sequences were tabulated and analyzed variously to test the following specific statements.

(a) There is no difference among Speech Reception scores for the six conditions of response delay of zero to five seconds.

(b) There is no difference among Speaker Intelligibility scores when the talker who receives his verbal material aurally delays his repetitions by zero to five seconds.

(c) There is no difference between write-down Speech Reception scores when one group of listeners responds after a one-second delay and another group responds after a five-second delay to the stimulus words.

(d) There is no difference between the Speaker Intelligibility scores for PB write-down words when the two response delay times are compared.

(e) There is no difference between the multiple-choice and PB reception scores or between the two types of speaker intelligibility scores when the two types of speech testing are compared for the one- and the five-second delay response times.

MULTIPLE CHOICE TESTS

(a) Speech Reception

The mean panel scores for each delay time were tabulated and arrayed for an analysis of variance (Lindquist, Treatments by Subjects Design). A summary of the analysis is found in Table I following.

The analysis of variance indicates significant differences among mean panel scores for the delayed response times. As is apparent, there were differences among groups of listeners. The averages of the mean panel scores are found in Table II and are plotted in Figure 1. Computations for t ratios yield valı indicating that a difference between means of 2.12 percentage points is significant at the 1 per cent level of confidence; for 5 per cent the required difference is 1.61.

Table I

Summary of an Analysis of Variance of the Mean Correct Listener Reception Scores for Each of Six Different Response Delay Times (0-5 seconds). N panels = 12

Source of Variance	df	Sum of Squares	Variance	F
Delay (D)	5	234.21	46.842	11.356
Groups (G)	11	4869.78	442.707	
D x G	55	226.89	4.125	
Total	71	5330.88		

$F = ms_D/ms_{D \times G}$, significant at the 1 per cent level of confidence, 5 and 55 df

Table II

Mean Panel Per Cent Correct Reception Scores at Each of Six Response Delay Times. N = 12

Response Delay Times (sec.)	0	1	2	3	4	5
Per Cent Correct	81.6	81.1	79.3	82.5	82.7	85.2

It is apparent from Table II and Figure 1 that there was a general increase in the mean reception scores with increased response delay times over the range explored, with the exception of the two-second time. Even including the decrease at two seconds the function appears to be essentially linear with the five-second delay time yielding the highest speech reception scores.

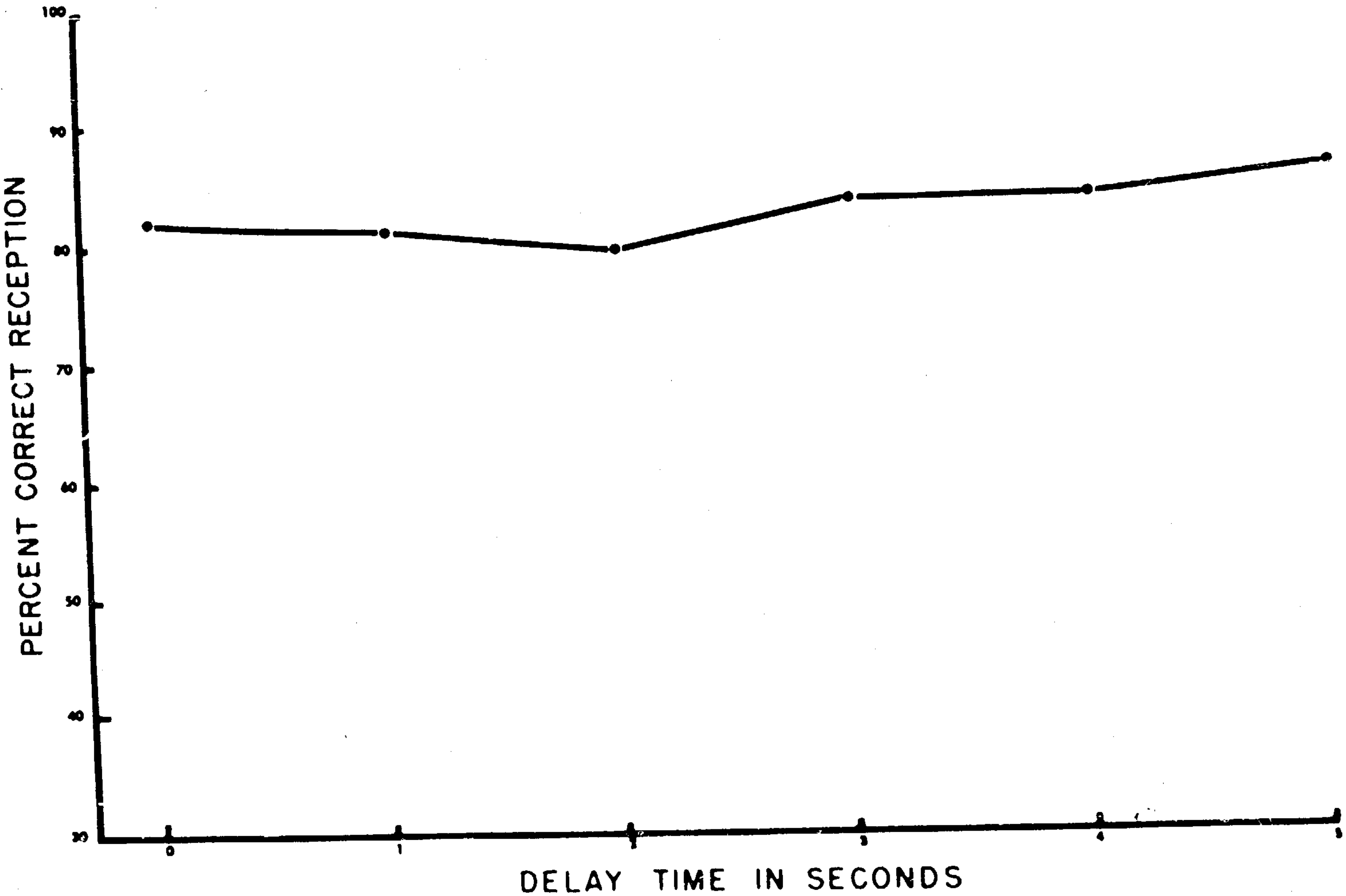

FIGURE I

MEAN PANEL RECEPTION SCORES FOR MULTIPLE-CHOICE TESTING WHEN THE LISTENER RESPONSES ARE DELAYED 0 TO 5 SECONDS

(b) Speaker Intelligibility

The mean speaker intelligibility scores for each of the 48 talkers at each of the six response delay times were computed and patterened to form the basic measures for an analysis of variance (Lindquist, Treatments by Subjects design). A summary of the analysis is given in Table III. The mean intelligibility scores for each speaker were contributed by 16 to 30 listeners.

Table III

Summary of an Analysis of Variance of the Mean Per Cent Correct Speaker Intelligibility Scores at Each of Six Response Delay Times (0-5 seconds). N = 48

Source of Variance	df	Sum of Squares	Variance	F
Delay (D)	5	3324.71	664.94	11.129*
Speakers (S)	47	9362.31	199.20	3.334
D x S	235	14041.19	59.75	
Total	287	26728.21		

*$F = ms_D / ms_{D \times S}$, significant at the 1 per cent level of confidence; 5 and 235 df.

The above analysis shows significant differences among the response delay times as these affect the speaker intelligibility scores of talkers who delayed their repetitions of aurally presented verbal material. The variance for speakers was also significant, as would be expected.

The data obtained from the 48 speakers were arrayed so as to yield mean panel speaker intelligibility scores. These means were composed of the averages per condition of the four speakers who were heard by a common panel of listeners. An analysis of variance (Treatments by Subjects) was made on the above data to confirm the means and indicate group homogeneity. The F ratio among response delay times was 12.063, 5 and 55 df, significant at the 1 per cent level of confidence. The F ratio for speaker groups was nonsignificant.

The mean speaker intelligibility scores for both tabulations were identical for all conditions of response delay times. These are listed below in Table IV and are plotted in Figure 2. The t-ratio computations indicate that a difference between means of 4.05 is significant at the 1 per cent level of confidence, and a difference of 3.08 is significant at the 5 per cent level.

Table IV

Mean Speaker Intelligibility Per Cent Correct Scores for Each of the Six Conditions of Response Delay Times. N 48

Response Delay Times (sec.)	0	1	2	3	4	5
Per Cent Correct	76.4	71.8	74.0	78.7	75.0	82.4

An examination of the means in Table IV and the plot of those means in Figure 2 reveals a general increase in speaker intelligibility scores as the response delay time increased, particularly if the zero time delay is ignored. It would also appear that the function seems to be essentially linear with the highest intelligibility scores resulting from the five-second delay time. This parallels the results obtained from the multiple-choice speech reception portion of the study.

PB WORD TESTS

(c) Speech Reception

The PB reception scores of the two groups of listeners, one group responding with a one-second delay the other a five-second delay, were analyzed for differences by a t-for related measures. The value of the t-ratio was 12.59, which is significant at the 1 per cent level of confidence indicating a difference in reception write-down scores between the one- and the five- seconds response delay times to the one-voiced recording. The mean panel scores for the two delay times were 91.1 per cent for the one-second condition and 84.3 per cent for the five-second delay, a reversal from the results obtained from the multiple-choice testing circumstance.

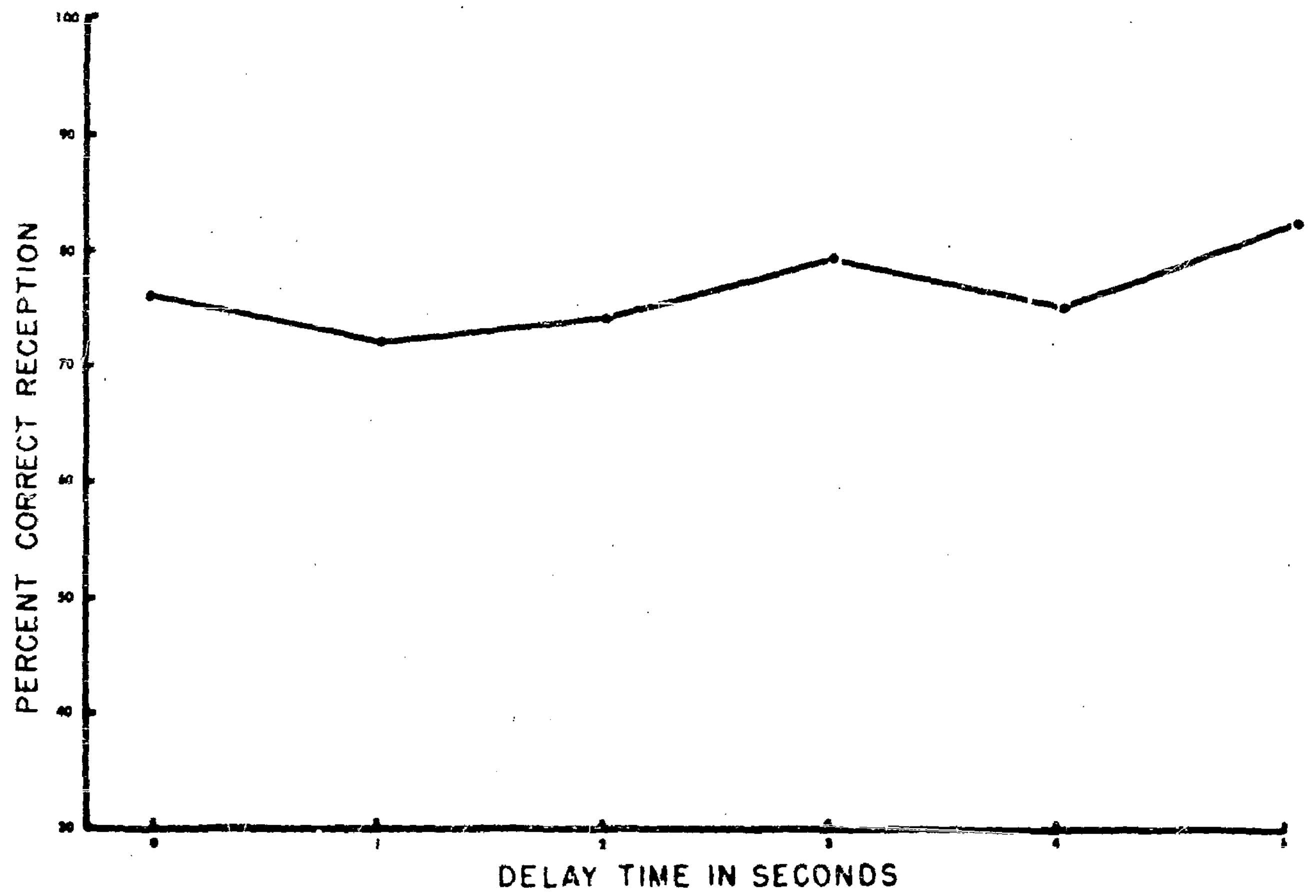

FIGURE 2

MEAN SPEAKER INTELLIGIBILITY SCORES OBTAINED FROM MULTIPLE-CHOICE INTELLIGIBILITY TESTS WHEN THE SPEAKER DELAYED SAYING THE THREE-WORD ITEM GROUPS HEARD AURALLY BY 0 TO 5 SECONDS.

(c) Speaker Intelligibility

Each of the 24 individuals who served as a speaker made the repetitions of the words he heard plus the carrier phrase under the one- and five-second response delay times, 50 words per speaker per condition. Two mean intelligibility scores were obtained for each individual, each condition. These scores were tabulated and analyzed by an analysis of variance (Treatment by Subjects design). The summary is found in Table V.

Table V

Summary of an Analysis of Variance of Mean Per Cent Correct Speaker Intelligibility Scores Obtained Under a One- and a Five-Second Response Delay Time. (PB Words.) N 24

Source of Variance	df	Sum of Squares	Variance	F
Delay (D)	1	202.95	202.95	4.077*
Speakers (S)	23	2098.55	91.24	1.833
D x S	23	1194.97	49.78	
Total	47	3446.47		

*$F = ms_D/ms_{D \times S}$, significant at the 6 per cent level of confidence, 1 and 23 df

The difference between means obtained under the conditions of one-second delay (82.3 per cent) and of five-second delay (86.4 per cent) is in the same general direction as for the multiple-choice speaker intelligibility situation reported above. The analysis shows, however, the differences just miss being significant. The differences between means are reversed from the PB word reception results.

A series of t's for unrelated measures were computed for a sample of 8 of the 24 speakers in which one-half of the scores assigned to a speaker for one-second delay were compared with one-half of the scores for five seconds-response-delay time. A second set of t-ratios were obtained on the other one-half of the scores, reversing the pattern of the first set. The values and their significance levels are found in Table VI.

Table VI

A Distribution of $\underline{t}$ (Unrelated) Values for Eight Speakers Comparing Split-Half Scores Between One- and Five-Seconds Response Delay Times. Levels of Significance are also Given. df = 48

Speakers	$\underline{t}$ First Set	% Significance	$\underline{t}$ Second Set	% Significance
1	23.92	1	24.94	1
2	25.88	1	34.34	1
3	8.67	1	5.35	1
4	7.20	1	13.00	1
5	.63	55	1.95	8
6	8.86	1	1.53	12
7	1.27	22	8.58	1
8	3.57	1	3.02	1

The values of $\underline{t}$ in Table VI show a high degree of consistancy for five of the eight speakers when the differences between response delay times are compared. Three of the eight show marked variability.

MULTIPLE-CHOICE AND PB WORD TESTING COMPARISONS

(a) Listener Reception

Two series of scores were taken from two listening panels of the multiple-choice and from all of the PB word reception data at each of the one- and five-second response delay times. These were arrayed to yield two $\underline{t}$-ratios for unrelated measures. The listening conditions and the sound-pressure levels of the signals were identical. The degree of listener sophistication was also the same.

At the one-second response delay time the mean panel score average for the multiple-choice reception was 78.6 per cent and 91.1 per cent for the PB words. The $\underline{t}$-ratio was 6.07 indicating significant differences between the means of the listener's scores for the two testing methods. The five-second delay yielded means of 83.9 per cent for the multiple-choice and 84.3 per cent for the PB words. In this instance the $\underline{t}$-ratio was 0.278, nonsignificant.

(b) Speaker Intelligibility

The speaker intelligibility scores from 48 individuals repeating the multiple-choice word phrases were compared with the scores of 24 individuals repeating the PB words under the one-and five-second response delay times. The t-ratio for the one-second delay was 4.63, showing differences significant at the 1 per cent level of confidence. The mean per cent correct scores were: multiple choice, 71.8 and for the PB, 86.4.

For the five-second response delay time the t-ratio for unrelated measures was 2.10, revealing differences among means significant at the 5 per cent level of confidence. The mean per cent correct scores were: multiple-choice, 82.3 and for the PB, 86.4.

DISCUSSION

The results of the speech reception and speaker intelligibility functions when the multiple-choice tests were used as the stimulus materials at each of six randomized conditions of response delay times follow similar trends that seem to indicate that as the delay time is increased between the stimulus word groups and either speaker or listener responses, reception and/or speaker intelligibility is enhanced. This was not completely in accord with casual observation or reports of listeners or speakers when questioned concerning the certainty of their responses.

However, in retrospect, the results seem logical, particularly with respect to listener reception of multiple-choice test items. One fundamental premise of the test is that all of the informational content of the speech sample is before the individual at all times. Giving the listener additional time may allow him to scan the alternative words sub-vocally or soto voce, thus making more accurate the "comparisons" with remembered acoustic patterns.

The possible explanation above could hardly be valid when attempting to account for the results obtained from the multiple-choice speaker intelligibility data. The talker did not have the alternative word choices before him as he repeated the word groups he heard. However, the subjective observations of experimenters and of some of the experimental subjects report a "process" of sub-vocal repetitions of each phrase, over and over, until the tonal signal to commence talking was received. At this instant the individual tended to "burst" into speech with the verbal material well rehearsed. It is true that he could erroneously rehearse mishearings, but the evidence seems to indicate that increasing the time of such a sub-vocal repetitive process enhanced intelligibility. Of course, there is no direct evidence that such a process was operating.

The results certainly make evident that an extension of the zero to five seconds response delay times should be explored experimentally in order to determine at which delay times the trend found in the present study would be reversed.

The data concerning the two response delay times when the speech material was the PB word lists present a more equivocal picture. Whereas the speaker intelligibility portion of the study exhibits the same general trends as in the multiple-choice testing circumstance, the PB word listener reception data show a reversal, in that higher reception scores were obtained under the condition of one-second response delay time.

It is possible that when the talker was repeating the word he heard, the highly theoretical process of increased rehearsal time "solidifying" and enhancing his vocal responses could be operating. If we assume that "rehearsing" is in progress here, an additional step of abstraction could be predicated that perhaps a small portion, one syllable of verbage, makes for unfavorable rehearsal material when compared with larger portions of from four to five syllables. This then might be an explanation for the higher speaker scores being associated with the longer response delay time.

One possible assumption that could be made to "explain" the results of the reversed speech reception data is that some of the increase rehearsal time confusedly could be spent in concern over spelling and legibility. A tenuous explanation, it is true. Additional experimentation is indicated with PB word lists using all of the response delay times employed in the multiple-choice testing portion of the present study and using the same extensions of delay time proposed above.

The comparisons between the multiple-choice test results and the PB test results were made primarily as a "finger-exercise" yet with the hope that some insight might be gained with respect as to how both could be employed in similar situations. Statistical differences were found among three of the four comparisons.

REFERENCES

1. Atkinson, C. J., Adaptation to delayed side-tone 1. Joint Project NM 18 02 99, Subtask 1, Report No. 12. Pensacola, Fla.: Ohio State Univ. Res. Found. and Naval School of Aviation Medicine, 1952.

2. Black, J. W., Tolhurst, G. C., and Morrill, S.N., Applications of multiple-choice speech intelligibility tests in the evaluation and use of voice communication equipment. Joint Project NM 18 02 99, Subtask 1, Report No. 19. Pensacola, Fla.: Ohio State Univ. Res. Found. and Naval School of Aviation Medicine, 1953.

3. Egan, J. B., Articulation Testing Methods, II. OSRD Report No. 3802. Cambridge, Mass.: Harvard Psycho-Acoustic Laboratory, 1945.

4. Peters, R. W., Competing messages: The effect of interfering messages upon the reception of primary messages. Joint Project NM 18 02 99, Subtask 1, Report No. 27. Pensacola, Fla.: Ohio State Univ. Res. Found. and Naval School of Aviation Medicine, 1954.

5. Peters, R. W., Listener performance as a function of listening time for various signal-to-noise conditions. Joint Project NM 18 02 99, Subtask 1, Report No. 53. Pensacola, Fla.: Ohio State Univ. Res. Found. and Naval School of Aviation Medicine, 1955.

6. Peters, R. W., The effect of recall and problem-solving tasks, concurrently presented with intelligibility testing materials, upon listener responses to the intelligibility materials. Joint Project NM 18 02 99, Subtask 1, Report No. 5?. Pensacola, Fla.: Ohio State Univ. Res. Found. and the Naval School of Aviation Medicine, 1955.

7. Tolhurst, G. C., Speech reception and hearing loss as a function to exposure to high-level noise. J. Acoust. Soc. Amer., 28: 557-560, 1956.

8. Tolhurst, G. C., The effect on intelligibility scores of specific instructions regarding talking. Joint Project NM 18 02 99, Subtask 1, Report No. 35. Pensacola, Fla.: Ohio State Univ. Res. Found. and Naval School of Aviation Medicine, 1954.

9. Tolhurst, G. C., Some effects of changing time patterns and articulation upon intelligibility and word reception, J. Speech Hearing Dis., In press, Sept. 1957.

10. Tolhurst, G. C., Speaker intelligibility of repeated messages acquired by visual, aural, or visual-aural channels. Joint Project NM 19 02 99, Subtask 1, Report No. 43. Pensacola, Fla.: Ohio State Univ. Res. Found. and Naval School of Aviation Medicine, 1955.

11. Tolhurst, G. C., The effects of an instruction to be intelligible upon a speaker's intelligibility, sound pressure level, and message duration. Joint Project NM 18 02 99, Subtask 1, Report No. 58. Pensacola, Fla.: Ohio State Univ. Res. Found. and Naval School of Aviation Medicine, 1955.

12. Tolhurst, G. C., The effects of disrupting the simultaneity of visual-aural communication channels to a speaker. Joint Project NM 18 02 99, Subtask 1, Report No. 66. Pensacola, Fla.: Ohio State Univ. Res. Found. and Naval School of Aviation Medicine 1956.

UNCLASSIFIED

AD 154613

Armed Services Technic

ARLINGTON HAL

ARLINGTON

FOR

MICRO CARD

CONTROL ONLY

1 OF 1

NOTICE: WHEN GOVERNMENT OR OTHER DRAWI
ARE USED FOR ANY PURPOSE OTHER THAN IN CO
GOVERNMENT PROCUREMENT OPERATION, THE
NO RESPONSIBILITY, NOR ANY OBLIGATION WHAT
GOVERNMENT MAY HAVE FORMULATED, FURNISH
SAID DRAWINGS, SPECIFICATIONS, OR OTHER DAT
IMPLICATION OR OTHERWISE AS IN ANY MANNER
PERSON OR CORPORATION, OR CONVEYING ANY R
USE OR SELL ANY PATENTED INVENTION THAT M

UNCLASSIFIED